DOCUMENT RESUME

ED 066 117 HE 003 281

AUTHOR	Cole, Nancy S.
TITLE	Bias in Selection.
INSTITUTION	American Coll. Testing Program, Iowa City, Iowa. Research and Development Div.
REPORT NO	ACT-RR-51
PUB DATE	May 72
NOTE	23p.
AVAILABLE FROM	Publication and Information Services Division, The American College Testing Program, P. O. Box 168, Iowa City, Iowa 52240 ($1.00)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Admission (School); *Civil Rights; *Discriminatory Attitudes (Social); *Equal Opportunities (Jobs); *Higher Education; Sex Discrimination; Social Discrimination

ABSTRACT

Possible bias in selection procedures used for employment and college admissions is of great social and educational importance. However, there are many different definitions of what constitutes bias, with each definition based on different values and with different implications for how selection should be accomplished. A number of these definitions of bias and their implications are examined in this document, and a new model of fairness based on equal opportunity for potentially successful applicants is presented. The equal opportunity model is suggested as an intuitively appealing and socially desirable model for use in many selection situations. (Author/HS)

ED 066117

ACT RESEARCH REPORT

No. 51

51

May 1972

BIAS IN SELECTION

Nancy S. Cole

PUBLISHED BY THE RESEARCH AND DEVELOPMENT DIVISION

THE AMERICAN COLLEGE TESTING PROGRAM

P. O. BOX 168, IOWA CITY, IOWA 52240

HE003281

BIAS IN SELECTION

ABSTRACT

Possible bias in selection procedures used for employment and college admissions is of great social and educational importance. However, there are many different definitions of what constitutes bias, with each definition based on different values and with different implications for how selection should be accomplished. A number of these definitions of bias and their implications are examined, and a new model of fairness based on equal opportunity for potentially successful applicants is presented. The equal opportunity model is suggested as an intuitively appealing and socially desirable model for use in many selection situations.

Prepared by the Research and Development Division
The American College Testing Program

For additional copies write:

Publication and Information Services Division
The American College Testing Program
P.O. Box 168, Iowa City, Iowa 52240

(Check or money order must accompany request.) Price: $1.00

BIAS IN SELECTION

Nancy S. Cole[1]

The issue of bias in the use of tests for the selection of minority group members for employment and for admission to college has received much attention in recent years. However, in spite of wide concern with the issue and its broad implications for our society, there has been little agreement about what constitutes bias or what procedures should be followed to alleviate it.

It is the purpose of this paper to identify the values and beliefs about fairness which are the bases for several definitions of bias and to provide actual procedures for the practitioner to follow to alleviate bias according to the definition he chooses. In addition, a new definition of bias based on the concept of equal opportunity for the potentially successful applicant regardless of group membership is presented and suggested as an intuitively appealing and socially desirable idea of fairness for many selection situations in employment and college admissions.

Definitions of Selection Bias

Six models of selection bias, or its converse selection fairness, will be considered. They will be referred to here as the quota model, the regression model, the Darlington model, the employer's model, the Thorndike model, and the equal opportunity model.

The Quota Model

The quota model of bias involves the idea that fairness lies in some specified proportional representation. For example, a procedure which requires a priori that half of those selected must be men and half women is based on a quota model. Similarly, another quota model might require that the proportion of minority members employed by a firm match the proportion of minority members in the population. In both cases, the proportional representation of particular groups is specified a priori on the basis of value judgments about fairness, and any procedure which fails to yield the specified proportions is considered biased.

The Regression Model

The regression model of test bias follows from definitions of bias which deal with consistent errors of prediction. For example, Cleary (1968) defined bias in the following way:

> A test is biased for members of a subgroup of the population if, in the prediction of a criterion for which the test is designed, consistent nonzero errors of prediction are made for members of the subgroup [p. 115].

[1]The author acknowledges the many helpful suggestions of James W. L. Cole, Gary R. Hanson, Leo A. Munday, and Melvin R. Novick in the preparation of this paper.

Anastasi (1968) gave a similar definition: "Test bias refers to overprediction or underprediction of criterion measures [p. 559]." These definitions assume that fairness is achieved by selecting on the basis of predictions of a criterion score and lead to the examination of regression equations in the separate groups for consistent errors of prediction.

The regression model of bias has been followed in a number of empirical studies of bias in the use of tests in college admissions (e.g., Bowers, 1970; Cleary, 1968; Temp, 1971). These studies have been concerned with possible bias in the use with minority group members of regression equations based on a majority group. If the regression lines are identical in the groups, then the use of a single prediction equation is considered fair. However, if the equations are not identical, then separate regression equations must be used according to this definition of fairness.

Under a "fair" regression procedure in which separate within-group regression equations are used, the selecting institution is assured of the selection of those applicants with the highest predicted criterion scores on the basis of the available predictor variables. However, if the prediction is poorer in one group than in another, then the selection cutoff point will be relatively higher in the group with the poorer prediction. Intuitively, when prediction is poor, one might wish the cutoff points to be lowered to reflect the increased uncertainty. Under the regression model the opposite occurs, and members of groups for which prediction is poor are penalized in the selection process.

The Darlington Model

Darlington (1971) argued that fairness can be achieved only by a kind of combination of the regression model and the type of value judgments made in the quota model. According to this model, one must first decide if there is special value in the selection of members of some cultural group. If so, then one accepts some difference between criterion scores which will yield equally desirable candidates from different groups.

For example, if it is valuable to obtain minority group members, one might decide that a minority member's score of Y on a criterion is as desirable as a score of Y + k on the criterion for majority members. Using a variable C which has value zero for minority group members and one for majority group members, a difference of k on Y is equivalent to a difference of one on C. Thus, by selecting on the basis of the variable Y − kC, the subjective judgment about the importance of selecting minority group members can be implemented; and in Darlington's terms a culturally optimum procedure is achieved. When k is set equal to zero (when there is no reason to favor one cultural group), this model reduces to the regression model.

The Employer's Model

Another definition of bias has led to a different selection model. Guion (1966) stated that

> unfair discrimination exists when persons with equal probabilities of success on the job have unequal probabilities of being hired for the job [p. 26].

This definition was implemented in a model proposed by Einhorn and Bass (1971).

Einhorn and Bass, by considering the distribution of criterion scores about the regression line, prescribed predictor cutoff points for each subgroup above which applicants have a specific minimal chance of being successful (or scoring above some specified criterion). For example, suppose an employer (or selector) is willing to hire all applicants with at least a 70% chance of success (or a 30% risk) as gauged by the predictor variables used. Then the predictor cutoff is chosen at the point at which the criterion pass point (Y_p) is approximately one-half standard error of estimate ($\sigma_{y \cdot x}$) below the predicted criterion ($\hat{Y}$) since about 70% of the cases fall above minus one-half standard deviation in a normal distribution. In terms of a unit normal deviate Z_p, where $Z_p = (Y_p - \hat{Y})/\sigma_{y \cdot x}$, $\Pr\{Z > Z_p\} = .70$.

Because the employer (or selector) can set the level of risk he is willing to assume, this model is especially advantageous to the employer—hence,

the reference to it as the employer's model. However, as with the regression model, poor prediction in one group decreases the chances of selection of members of that group. When the prediction is poor, the standard error of estimate is large; consequently, a higher predicted score (and predictor cutoff) is required to maintain the same one-half standard error of estimate difference between predicted criterion and the criterion pass point in the example. Thus, poor prediction lowers the chances of success of a person with a high predictor score and consequently decreases his chance of selection.

The Thorndike Model

Thorndike (1971) proposed yet a fifth definition of bias or its complement, fairness. In a fair selection procedure,

> the qualifying scores on a test should be set at levels that will qualify applicants in the two groups in proportion to the fraction of the two groups reaching a specified level of criterion performance [p. 63].

Although more complicated sounding at first glance, Thorndike's is actually a very simple notion. If, in Group A, 50% of the members are successful and, in Group B, 80% of the members are successful, then the proportion of Group A members selected to those selected from Group B should match the 50:80 success ratio. Thus, Thorndike's model requires that the success ratio equal the selection ratio, or in terms of probability statements,

$$\frac{\Pr_1\{Y > Y_p\}}{\Pr_2\{Y > Y_p\}} = \frac{\Pr_1\{X > X_1\}}{\Pr_2\{X > X_2\}} \qquad (1)$$

where X is the predictor variable in the two groups, Y the criterion, Y_p the criterion "pass point" or the predetermined criterion level of success, and X the selection cutoff points in the two groups on the predictor variable.

Thorndike's idea of fairness as a match of selection rate to success rate has intuitive appeal in that it eliminates the inequity of over-selecting in a group in which prediction is better even when a substantial proportion of the group with poorer prediction could succeed if selected. Thus, whereas the regression and employer's models are advantageous primarily from the selecting institution's point of view, Thorndike's model proposes a kind of fairness more nearly appropriate from the applicant's viewpoint. The model of bias proposed next can be seen as a logical extension and refinement of the Thorndike model to an even more intuitively appealing idea of fairness to the applicant in a selection process.

The Equal Opportunity Model

In many selection situations, the applicant who, if selected, would be able to succeed deserves a guarantee of fairness in selection. Usually, not all potentially successful applicants can be selected both because too few positions are available and because one is unable to identify in advance with surety who will and who will not succeed. However, when the distribution of a predictor and a criterion of success are known by past experience, one can compute the probability that a potentially successful applicant has of being selected given a fixed selection procedure.

Under each of the previous models discussed, it may happen that the chance of selection of a potentially successful applicant in Group A is different from the chance of selection of such an applicant in Group B. Thus, two applicants, both of whom could succeed (achieve a criterion score above a criterion pass point) if selected, may have different chances of selection because of their group membership. Under the equal opportunity model this type of unfairness is eliminated.

The principle of the equal opportunity model is that, as a group, people who can achieve a satisfactory criterion score ($Y > Y_p$) should have the same probability of being selected whether

minority or majority group members. In terms of probability statements, the equal opportunity model specifies that a selection procedure is fair when

$$\Pr_1\{X > X_1 \mid Y > Y_p\} = \Pr_2\{X > X_2 \mid Y > Y_p\}. \quad (2)$$

Thus, equal opportunity as defined in this model is equal opportunity to those who could be successful. If a predictor cutoff is set in one group so that the probability of being selected when potentially successful is .80, then the model requires that to be fair the predictor cutoff must be set in the other subgroup to give the same conditional probability.[2]

The equal opportunity model, like the employer's model and Thorndike's model, requires the specification of a criterion pass point (Y_p) above which performance is satisfactory and below which, unsatisfactory. Although it is probably reasonable to set such a point in most situations of employment and college admissions, at the same time the selecting institution often is concerned with degrees of relative success. The equal opportunity model, like the other two models, uses a zero-one utility model in which utility for degrees of relative success is not included. However, when a selecting institution rewards with graduation or continuing employment those who achieve a minimal level of competence, the zero-one utility model is certainly relevant to the selection process.

Applications of the Bias Models to a Selection Situation

Six definitions of bias have been presented, and each has some point of intuitive appeal. However, there are many situations in which they yield quite different answers to the question, "Is a selection procedure biased?" In this section a type of selection situation is described and the prescriptions for fairness which each definition of bias yields are derived in order to provide a common ground for direct comparison of the models of bias in several hypothetical situations.

A Selection Situation

In this paper a selection situation will involve a certain number of applicants N_i from each of several groups and a number of available openings N_o where $N_o < \sum_i N_i$. For simplicity only two groups in this section and in the examples which follow in the next section are considered. Also only the case of a single predictor is considered although the results are identical for multiple prediction when the within-group multiple regression equation is used as a single predictor.

Selection is accomplished through the use of a predictor variable[3] X which in each group has a known relationship to a criterion Y. In the cases examined here, it is assumed that X and Y have a bivariate normal distribution in each group and that the means ($\mu_{x(i)}$ and $\mu_{y(i)}$), the standard deviations ($\sigma_{x(i)}$ and $\sigma_{y(i)}$), and the correlation ($r_{xy(i)}$) are known from past experience. Then, for a criterion pass point or success point, Y_p, the selection problem is to choose predictor cutoff points in each group, X_i, so that N_o applicants are selected and that the particular fairness model is satisfied.

---

[2]Darlington (1971) described four definitions of culturally fair tests in terms of the correlation, r_{xc}, of the predictor variable X and a cultural variable C and then rejected all of the definitions in favor of the Darlington model described above. However, it is interesting to note that the equal opportunity model satisfies Darlington's (p. 73) definition (3) which requires that $r_{xc} = r_{cy}r_{xy}$; and consequently the present model provides an entirely different rationale for that definition.

[3]It is possible to use different predictors for the different groups under each model except Darlington's model. However, for notational simplicity the same predictor variable X is dealt with in each group.

In order to achieve N_O selectees, the following equation must be satisfied.

$$N_1 Pr_1\{X > X_1\} + N_2 Pr_2\{X > X_2\} = N_O. \quad (3)$$

Thus, by simultaneously satisfying (3) and its own restrictions of fairness, each model will specify the values of X_1 and X_2.

The Bias Models as Selection Models

An extension must be made in each of the models of bias in order to solve for the predictor cutoffs X_1 and X_2 in the selection procedure described above.

In the quota model, for each group the proportion of the total selected, the quota, is set at, say, p_i where

$$p_i = \frac{N_i Pr_i\ \{X > X_i\}}{N_O} \quad (4)$$

Thus, in each group, X_i is set so that (4) is satisfied. If there are 100 applicants from group i and 50 openings and the group's quota is 50% of those selected, then X_i must be set so that $Pr_i\{X > X_i\} = .25$.

Under the regression model, the separate within-group regression lines are used to select the N_O applicants with the highest predicted criterion, $\hat{Y}$. Therefore at that predicted criterion cutoff point, $\hat{Y} = a_1 + b_1X_1 = a_2 + b_2X_2$. Thus, X_2 can be expressed in terms of X_1 as follows:

$$X_2 = (a_1 - a_2 + b_1X_1)/b_2 \quad (5)$$

Then, for any value of X_1, X_2 can be computed, and $Pr_1\{X > X_1\}$ and $Pr_2\{X > X_2\}$ can be read from tables of the normal distribution. By substituting $Pr_i\{X > X_i\}$ into (3), the number of applicants which would be selected using these values of X_1 and X_2 can be computed. If that number does not equal N_O, a new value of X_1 must be selected and the iterative process continues until values of X_1 and X_2 which yield precisely N_O selectees are found.

Under the Darlington model, the data from the two groups are combined, and a prediction equation for Y is computed using both X and C, where C is a dichotomous variable in which C = 0 for Group 1 (the minority group) and C = 1 for Group 2 (the majority group). Then kC is subtracted from that equation to give the prediction equation for $\widehat{Y - kC}$ (Darlington, 1971, p. 81). If $\widehat{Y - kC} = c + dX + eC$ is the resulting equation then,

$$c + dX_1 = c + dX_2 + e \quad (6)$$

and

$$X_2 = X_1 - e/d \quad (7)$$

Thus, as was discussed with the regression model, using an iterative procedure a pair of probabilities $Pr_i\{X > X_i\}$ in which X_1 and X_2 are related as in (7) and which satisfy (3) can be found.

In the employer's model as described by Einhorn and Bass (1971), the first step is the specification of the risk the employer is willing to take (or the minimal chance of success he will allow in a selectee). Rather than setting the risk a priori one assumes that the employer wants to fill N_O openings and is willing to adjust the risk level to get them, so long as the risk is the same in the two groups. Therefore, in this specification of the selection process, one looks for predictor cutoff points which will fill the N_O openings while at the same time keeping the employer's risk the same in both groups.

In equation (4) of Einhorn and Bass (1971, p. 266), one specifies the risk by choosing a unit normal deviate Z_p. The probability of a deviate above that value corresponds to the minimum tolerable chance of success. Rather than specifying Z_p, one requires only that $Z_p = (Y_p - \hat{Y})/\sigma_{y \cdot x}$ be

the same in each group. Again using $\hat{Y} = a_i + b_i X_i$, one can solve for X_2 in terms of X_1 as before.

$$X_2 = \frac{\sigma_{y\cdot x(2)} (a_1 + b_1 X_1 - Y_p)}{b_2\, \sigma_{y\cdot x(1)}} + \frac{Y_p - a_2}{b_2} \qquad (8)$$

where $\sigma_{y\cdot x(i)} = \sigma_{y(i)} \sqrt{1 - r^2_{xy(i)}}$. Then, by finding by iteration the pair of probabilities $Pr_i\{X > X_i\}$ which satisfy (8) and (3), the required cutoff points X_1 and X_2 can be found.

With Thorndike's model, one solves (3) for $Pr_1\{X > X_1\}$ in terms of $Pr_2\{X > X_2\}$ and substitutes into (1) to obtain

$$Pr_2\{X > X_2\} = \frac{N_0}{N_1 \left[\frac{Pr_1\{Y > Y_p\}}{Pr_2\{Y > Y_p\}}\right] + N_2} \qquad (9)$$

$Pr_1\{X > X_1\}$ follows from (1) and the two probabilities imply values of X_1 and X_2 which are the required solution for the Thorndike model.

In the equal opportunity model, equations (2) and (3) must be simultaneously solved. The solution can be accomplished by expressing the conditional probabilities in (2) in terms of joint probabilities by the well-known relationship,

$$Pr_i\{X > X_i \mid Y > Y_p\} = \frac{Pr_i\{X > X_i,\ Y > Y_p\}}{Pr_i\{Y > Y_p\}} \qquad (10)$$

One first chooses a value of X_1 and using Y_p and the means and variances of X and Y and the correlation of X and Y, the value of $Pr_1\{X > X_1,\ Y > Y_p\}$ can be read from tables of the bivariate normal distribution (National Bureau of Standards, 1959). Then,

$$Pr_2\{X > X_2,\ Y > Y_p\} = \left[\frac{Pr_2\{Y > Y_p\}}{Pr_1\{Y > Y_p\}}\right] Pr_1\{X > X_1,\ Y > Y_p\} \qquad (11)$$

can be computed using $Pr_1\{X > X_1, Y > Y_p\}$ and $Pr_i\{Y > Y_p\}$ from tables of the normal distribution. From the bivariate normal tables the value of X_2 yielding a particular $Pr_2\{X > X_2, Y > Y_p\}$ can be found. Then for X_1 and X_2, $Pr_1\{X > X_1\}$ and $Pr_2\{X > X_2\}$ can be computed. The iteration continues until values of $Pr_i\{X > X_i\}$ which satisfy (3) are found.

The actual solutions the models produce in several hypothetical situations will now be considered. Only the last five models will be explicitly examined. However, the proportions of each group selected will be presented so that any quota model can be compared with the results of the other models.

Comparison of Selection Models

In many discussions of bias in college admissions and employment (e.g., Anastasi, 1968; Bartlett & O'Leary, 1969), the possibility that the minority regression line is parallel to but above the majority line has been of great concern since the majority regression equation is often used for selection in both groups. Thus, the comparison of the selection models is begun by considering this situation in Case A.

However, in available empirical studies involving minority racial/ethnic groups, especially in the area of college admissions, the type of situation just described is very rare. According to these studies (e.g., Bowers, 1970; Cleary, 1968; Temp, 1971), it is much more common for one of four situations to occur: (1) the regression lines for majority and minority groups may be quite similar, but with lower minority group means on both the predictor

and the criterion; (2) the slope of the minority regression line may be smaller than that for the majority group, or the two lines may differ both in slope and intercept with the majority line lying above the minority line in the region of practical predictions and (3) the minority line having the smaller slope or (4) the majority line having the smaller slope. Thus, these four situations will also be explored in Cases B, C, D, and E in order to compare the prescriptions for fairness of the various bias models in commonly found situations.

Case A

In Case A, which corresponds to case A of Thorndike (1971) and is specified in Figure 1, the two regression lines are parallel but with different intercepts. In each of the cases considered, the minority group is Group 1 and the majority, Group 2, and selection is required since there are 500 applicants (100 minority and 400 majority) for 100 openings.

Because the regression lines differ in Case A, according to the regression model the separate regression equations must be used to select the 100 students with the highest predicted grades. Using (5) we find that $X_2 = 1 + X_1$. Because the means also differ by one, the standard normal cutoff variable $Z_i = (X_i - \bar{X})/\sigma$ is the same in the two groups. Thus, $Pr_1\{Z > Z_1\} = Pr_2\{Z > Z_2\} = N_0/(N_1 + N_2)$, and one need only use standard normal tables to find the Z above which 20% $[N_0/(N_1 + N_2)]$ of the applicants fall. Since $Pr\{Z > 0.842\} = .20$, .842 is that cutoff. For Group 1 one must convert back to the original scores where the cutoff becomes —.158. Thus, according to the regression model, one should select all minority group members with test scores above —.158 and all majority group members above .842, as summarized in Table 1. This procedure will yield the required 100 persons and will be fair in the sense that the 100 persons with the highest predicted criterion scores will be chosen.

In the Darlington model the minority group is favored by predicting (Y —, .5C). The required

	Minority	Majority
$\mu_{x(i)}$	−1.0	0.0
$\sigma_{x(i)}$	1.0	1.0
$r_{xy(i)}$	0.5	0.5
$\mu_{y(i)}$	0.0	0.0
$\sigma_{y(i)}$	1.0	1.0

$Y_p = 0.0$

$N_1 = 100$, $N_2 = 400$, $N_0 = 100$

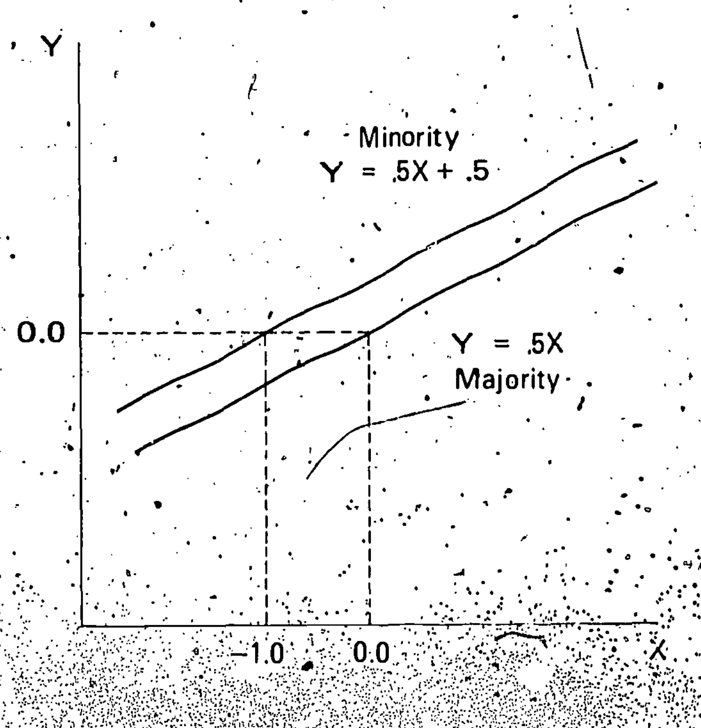

Fig. 1. Case A in which the regression lines are parallel with the minority line falling above the majority line.

TABLE 1

Results of Selection Models for Five Common Selection Situations

		Regression model	Darlington model	Employer's model	Thorndike model	Equal opportunity model
Case A—Parallel regression lines, minority intercept larger						
	Cutoff point	−.16	−.90	−.16	−.16	−.16
Minority:	Percentage selected	20.0%	46.0%	20.0%	20.0%	20.0%
	Number selected	20	46	20	20	20
	Cutoff point	.84	1.10	.84	.84	.84
Majority:	Percentage selected	20.0%	13.6%	20.0%	20.0%	20.0%
	Number selected	80	54	80	80	80
Case B—Identical regression lines, minority means smaller						
	Cutoff point	.71	−.16	.71	.11	−.02
Minority:	Percentage selected	4.5%	20.0%	4.5%	13.3%	16.4%
	Number selected	4	20	4	13	16
	Cutoff point	.71	.84	.71	.78	.81
Majority:	Percentage selected	23.9%	20.0%	23.9%	21.6%	20.9%
	Number selected	96	80	96	86	84
Case C—Different regression slopes, minority slope smaller						
	Cutoff point	2.39	.22	3.24	.84	.60
Minority:	Percentage selected	.8%	41.3%	.1%	20.0%	27.4%
	Number selected	1	41	0	20	27
	Cutoff point	.68	1.05	.67	.84	.91
Majority:	Percentage selected	24.8%	14.6%	25.0%	20.0%	18.1%
	Number selected	99	58	100	80	72
Case D—Different regression slopes and intercepts, minority slope smaller						
	Cutoff point	3.00	1.11	3.16	.95	.47
Minority:	Percentage selected	0.0%	5.4%	0.0%	7.4%	16.6%
	Number selected	0	5	0	7	17
	Cutoff point	.68	.72	.67	.73	.81
Majority:	Percentage selected	24.8%	23.7%	25.0%	23.2%	21.0%
	Number selected	99	95	100	93	84
Case E—Different regression slopes and intercepts, minority slope larger						
	Cutoff point	1.63	1.34	1.57	.95	.80
Minority:	Percentage selected	1.7%	3.3%	1.9%	7.4%	9.7%
	Number selected	2	3	2	7	10
	Cutoff point	0.69	.70	0.69	.73	.76
Majority:	Percentage selected	24.4%	24.1%	24.5%	23.2%	22.5%
	Number selected	97	96	98	93	90

prediction equation, following the procedure described by Darlington (1971, p. 81), is $\widehat{Y - .5C} = .5 + .5X - 1.0C$. By equation (7) one finds that $X_2 = X_1 + 2.0$. Solving iteratively one finds that predictor cutoff points of −.90 for the minority group and 1.10 for the majority group produce the desired 100 selectees. Thus, 46% of the minority group and 13.6% of the majority group would be selected by this procedure as given in Table 1.

In the case of the employer's model, equation (8) reduces to $X_2 = 1 + X_1$ which implies that $Pr_2\{Z > Z_2\} = Pr_1\{Z > Z_1\} = .20$, using (3). Consequently, as with the regression model $Z_1 = Z_2 = .842$ and in terms of raw scores $X_1 = -.158$ and $X_2 = .842$. This procedure is fair according to the employer's model because all those chosen have above a certain minimum probability of success. In Case A that probability can be computed by computing Z_p and can be shown to be .69. Thus, any person with a probability of success above .69, regardless of group membership, is selected.

In the Thorndike model since $Pr_1\{Y > Y_p\} = Pr_2\{Y > Y_p\}$, from (9) we see that $Pr_1\{X > X_1\} = Pr_2\{X > X_2\} = .20$; and again −.158 and .842 are the predictor cutoff points, respectively, required to achieve a fair procedure. Under this model the procedure is considered fair since the ratio of success probabilities, $Pr_1\{Y > Y_p\}/Pr_2\{Y > Y_p\} = 1.0$, is equal to the ratio of the selection probabilities, $Pr_1\{X > -.158\}/Pr_2\{X > .842\} = 1.0$.

For the equal opportunity model, when $r_{xy(1)} = r_{xy(2)}$ and $Pr_1\{Y > Y_p\} = Pr_2\{Y > Y_p\}$, then $Z_1 = Z_2$; and the same result as in the other three models is found here also. In this model, the result is considered fair because the potentially successful members of the minority group have a probability of .31 of being selected as do the potentially successful majority members.

Case B

In Case B the consideration of situations commonly found in empirical studies of racial/ethnic minorities is begun. Case B corresponds to Thorndike's (1971) case B and is presented in Figure 2. Because the regression lines are identical, according

	Minority	Majority
$\mu_{x(i)}$	−1.0	0.0
$\sigma_{x(i)}$	1.0	1.0
$r_{xy(i)}$	0.5	0.5
$\mu_{y(i)}$	−0.5	0.0
$\sigma_{y(i)}$	1.0	1.0

$Y_p = 0.0$

$N_1 = 100, N_2 = 400, N_0 = 100$

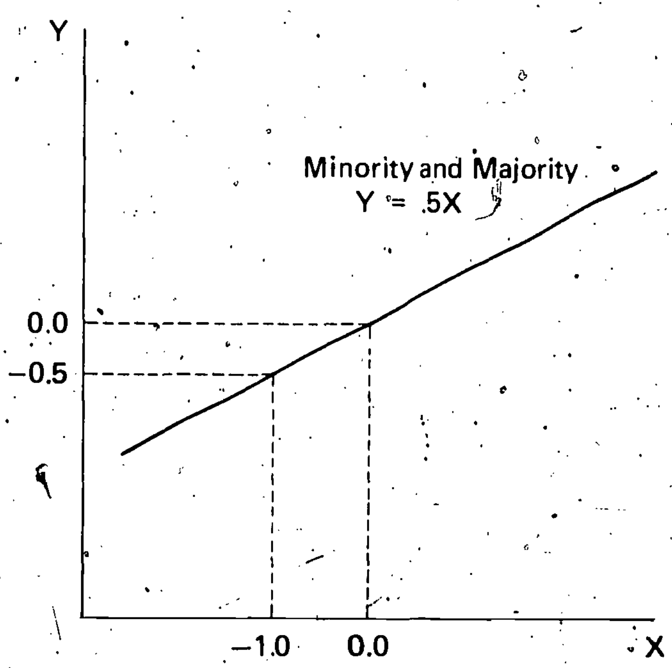

Fig. 2. Case B in which the regression lines are identical for the two groups but with lower minority group means.

to the regression model one needs select only the highest predicted grades or equivalently the highest test scores ($X_2 = X_1$ by (5)). A predictor cutting score of .71 in each group provides the 100 selections. Of course, since the minority group has lower test scores, only 4.5% of the minority group are selected while 23.9% of the majority applicants are accepted. The action is deemed fair because the same action is taken with all persons with equal criterion predicted scores regardless of group membership.

In the Darlington model, with k = .5 again in order to favor the minority group, the test cutoff points are −.16 and .84 for the minority and majority groups, respectively. Thus, 20% of each group would be selected by this procedure.

Under the employer's model, the cutoff points are .71 for the minority group and .71 for the majority group just as in the regression model. The same cutoffs for the groups here assure equal minimum probabilities of success among the selected members of both groups. In this case, the minimum chance of success of those selected is .66.

Since the success ratio of the minority to majority group is .617 in Case B, under the Thorndike model cutoff points of .11 and .78 are chosen because they provide 100 persons and a selection ratio of .617 as required. Thus, according to this idea of fairness, 13.3% of the minority group and 21.6% of the majority group are selected, giving considerably more minority group members than either the regression or employer's model.

Cutoff points of −.02 in the minority group and .81 in the majority group give 100 selectees and equal probabilities of selection for potentially successful applicants in both groups. Thus, these cutoff points are prescribed by the equal opportunity model. The computations required for the equal opportunity model in Case B are illustrated in Table 2. Under this model, still more minority group members are chosen, 16.4%; and in both groups $Pr_i\{X > X_i \mid Y > Y_p\} = .32$.

Thus, in Case B in spite of equal regression equations in the two groups, the different definitions of fairness call for dramatically different selection procedures as can be seen from the summary in Table 1. Many more minority students would be selected under Darlington's model, Thorndike's model, and the equal opportunity model than under the regression model and the employer's model.

Case C

In Case C the slopes of the regression of criterion on tests differ in the two groups as shown in Figure 3, and the basic results for each of the five models are reported in Table 1.

Very few minority students are selected in the regression and employer's models because of the small slope of the regression line in that group. When the slope is small, predicted scores are low as are the chances of a person with a high test score achieving a particular criterion score. Considerably more minority students would be selected under Darlington's model, Thorndike's model, and the equal opportunity model. In fact, when the means and variances of the groups are equal as in Case C, under the equal opportunity model a higher proportion of minority group than majority group members are selected precisely because of the small slope of the line in that group. If only a poor predictor is available, proportionately more applicants will have to be selected to insure that the potentially successful ones have the same opportunity for selection as members of a group with a larger regression slope.

Case D

In Case D the situation is examined in which both the slopes and the intercepts of the regression lines of the groups differ, with the majority line above the minority line in the region of practical interest and with a larger majority slope as shown in Figure 4. The results for each of the selection models are given in Table 1. In Case D no minority group members would be selected under the regression or employer's models even though 15.9% of the minority group could succeed if

TABLE 2

Computations for the Equal Opportunity Model in Case B

$\frac{Y_p - \mu_{y(1)}}{\sigma_{y(1)}} = 0.5$ $r_{xy(1)} = 0.5$		$\frac{Y_p - \mu_{y(2)}}{\sigma_{y(2)}} = 0.0$ $r_{xy(2)} = 0.5$				
Z_1	L_1	L_2	Z_2	Pr(1)	Pr(2)	N_o
0.00	.227	.368	−0.18	0.50	0.57	278
0.50	.163	.264	0.33	0.31	0.37	179
0.90	.110	.178	0.73	0.18	0.23	110
$\boxed{0.98}$	.100	.162	$\boxed{0.81}$	0.164	0.209	$\boxed{100.0}$
0.99	.098	.159	0.82	0.161	0.206	98.4

$Z_1 = 0.98 \rightarrow X_1 = -0.02$

$Z_2 = 0.81 \rightarrow X_2 = 0.81$

$$\frac{Pr_1\{X > 0.98, Y > 0.5\}}{Pr_1\{Y > 0.5\}} = \frac{.100}{.3085} = .324 = \frac{Pr_2\{X > 0.81, Y > 0.0\}}{Pr_2\{Y > 0.0\}}$$

Note. $L_1 = Pr_1\{X > X_1, Y > Y_p\}$ taken from National Bureau of Standards (1959, pp. 52-53).

$L_2 = [Pr_2\{Y > Y_p\}/Pr_1\{Y > Y_p\}]\ L_1 = (.5/.3085)L_1 = 1.62\ L_1$.

$Pr(i) = Pr_i\{X > X_i\}$.

$N_o = N_1Pr(1) + N_2Pr(2) = 100\ Pr(1) + 400\ Pr(2)$.

	Minority	Majority
$\mu_{x(i)}$	0.0	0.0
$\sigma_{x(i)}$	1.0	1.0
$r_{xy(i)}$	0.2	0.7
$\mu_{y(i)}$	0.0	0.0
$\sigma_{y(i)}$	1.0	1.0

$Y_p = 0.0$

$N_1 = 100, N_2 = 400, N_o = 100$

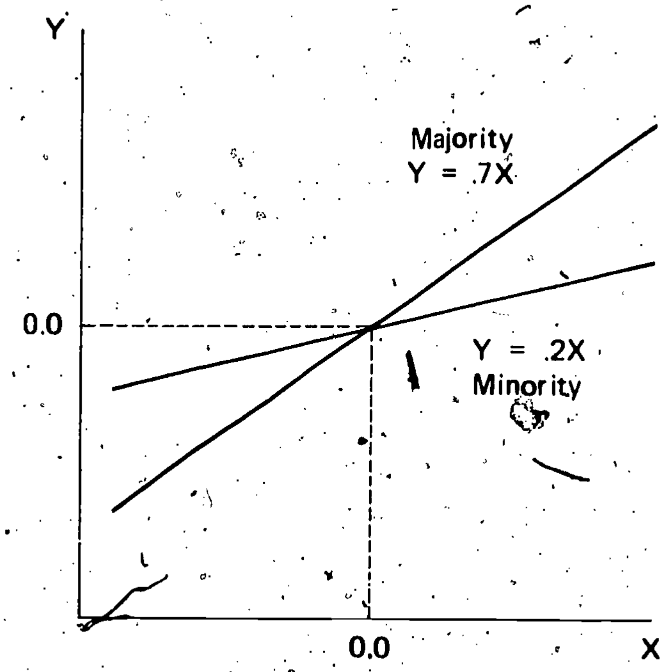

Fig. 3. Case C in which the two groups differ in the slope of the regression lines, with the minority slope smaller.

	Minority	Majority
$\mu_{x(i)}$	−0.5	0.0
$\sigma_{x(i)}$	1.0	1.0
$r_{xy(i)}$	0.4	0.6
$\mu_{y(i)}$	−1.0	0.0
$\sigma_{y(i)}$	1.0	1.0

$Y_p = 0.0$

$N_1 = 100, N_2 = 400, N_o = 100$

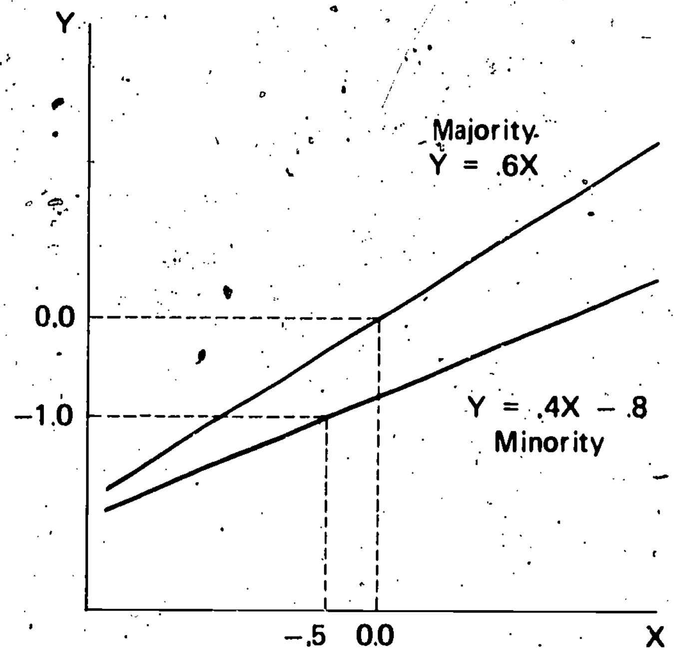

Fig. 4. Case D in which both the slopes and intercepts of the regression lines differ, with the minority slope smaller.

selected. By contrast, if equal opportunity for selection were granted the potentially successful minority members, 17% of them would be selected. One should note that even the Darlington model which explicitly favors the minority group on the predicted criterion score does not require the selection of as many minority applicants as required by the equal opportunity model.

Case E

Case E is similar to Case D except that in Case E the slope of the majority regression line is smaller than that of the minority group as shown in Figure 5. The increase in the minority slope results in the selection of two minority group members under the regression and employer's models compared with no selections in Case D, as can be seen in Table 1. However, the two models still do not provide equal opportunity to potentially successful minority members and neither does Darlington's model nor Thorndike's model as 10 minority members must be selected to provide equal opportunity. Thus, even in some situations in which prediction is better in the minority group than in the majority group but there are differences in the means, the regression and employer's models may be unfair to minority members according to the equal opportunity model.

Reversals of Minority and Majority Data in the Five Cases

Even though the cases examined are those of most importance in the considerations of bias against racial/ethnic minorities, the selection models apply to any identifiable groups in which other relationships of regression lines may occur. Consequently, the reversals of minority and majority data in the five cases are briefly examined here in order to illustrate the effects of the various models of bias in a wider variety of situations. The results for each of the models for the reversed data are given in Table 3.

	Minority	Majority
$\mu_{x(i)}$	−0.5	0.0
$\sigma_{x(i)}$	1.0	1.0
$r_{xy(i)}$	0.6	0.4
$\mu_{y(i)}$	−1.0	0.0
$\sigma_{y(i)}$	1.0	1.0

$Y_p = 0.0$

$N_1 = 100, N_2 = 400, N_0 = 100$

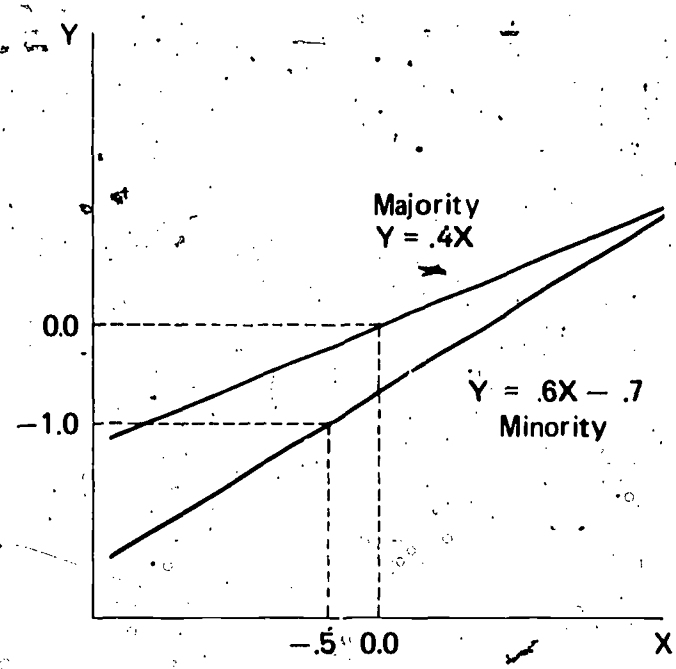

Fig. 5. Case E in which both the slopes and intercepts of the regression lines differ, with the minority slope larger.

TABLE 3

Results of Selection Models When Minority and Majority Data Are Reversed in the Five Cases

		Regression model	Darlington model	Employer's model	Thorndike model	Equal opportunity model
Case A Reversed—Parallel regression lines, minority intercept smaller						
	Cutoff point	.84	.10	.84	.84	.84
Minority:	Percentage selected	20.0%	46.0%	20.0%	20.0%	20.0%
	Number selected	20	46	20	20	20
	Cutoff point	−.16	.10	−.16	−.16	−.16
Majority:	Percentage selected	20.0%	13.6%	20.0%	20.0%	20.0%
	Number selected	80	54	80	80	80
Case B Reversed—Identical regression lines, minority means larger						
	Cutoff point	.10	−.55	.10	.56	.71
Minority:	Percentage selected	46.0%	70.9%	46.0%	28.8%	23.9%
	Number selected	46	71	46	29	24
	Cutoff point	.10	.45	.10	−.08	−.12
Majority:	Percentage selected	13.6%	7.4%	13.6%	17.8%	19.0%
	Number selected	54	29	54	71	76
Case C Reversed—Different regression slopes, minority slope larger						
	Cutoff point	.29	−.34	.22	.84	1.08
Minority:	Percentage selected	38.4%	63.3%	41.3%	20.0%	14.0%
	Number selected	38	63	41	20	14
	Cutoff point	1.02	1.33	1.06	.84	.79
Majority:	Percentage selected	15.4%	9.2%	14.5%	20.0%	21.3%
	Number selected	62	37	58	80	85
Case D Reversed—Different regression slopes and intercepts, minority slope larger						
	Cutoff point	−.65	−1.38	−.59	.15	.72
Minority:	Percentage selected	74.2%	91.6%	72.2%	44.1%	23.6%
	Number selected	74	92	72	44	24
	Cutoff point	1.025	1.53	.99	.62	.37
Majority:	Percentage selected	6.4%	2.1%	6.9%	14.0%	19.1%
	Number selected	25	8	27	56	76
Case E Reversed—Different slopes and intercepts, minority slope smaller						
	Cutoff point	−.43	−.97	−.46	.15	.40
Minority:	Percentage selected	66.6%	83.4%	67.7%	44.1%	34.5%
	Number selected	67	83	68	44	35
	Cutoff point	.88	1.23	.90	.62	.48
Majority:	Percentage selected	8.4%	4.2%	8.1%	14.0%	16.3%
	Number selected	34	17	32	56	65

When the data in Case A are reversed, for each of the models except Darlington's model, the cutoff points are simply reversed and, as before, 20% of each group is selected. The cutoff points under Darlington's model change because the combined group regression line is altered by the reversal, and even a larger number of minority applicants would be selected under that model than originally in Case A.

In the reversal of data in the remaining cases, even though the two cutoff points maintain the same relationship under the regression and employer's models as specified by equations (5) and (8), respectively, the particular pair of cutoff points which yield 100 selectees differs from each case to its reversal. Thus, in Case B the cutoff points of .71 in each group provide 100 selectees, but when the data are reversed, cutoff points of .10 in each group are required under both models to provide 100 selectees.

In the last four original cases (B through E), Thorndike's model and the equal opportunity model prescribed the selection of more minority group members in each case than the regression or employer's models. However, when the data are reversed, cases in which the regression and employer's models are unfair (by Thorndike's definition and the equal opportunity definition) to the majority group are illustrated.

Chances of Selection of Potentially Successful Applicants and Expected Success Rates

Table 4 gives the chances for selection which the potentially successful applicants have under each model in each of the five original cases. As is clear

TABLE 4

Potentially Successful Applicants' Chances of Selection

$Pr_i \{X > X_i \mid Y \geq Y_p\}$

	Regression model	Darlington model	Employer's model	Thorndike model	Equal opportunity model
Case A—Parallel regression lines, minority intercept larger					
Minority	.31	.63	.31	.31	.31
Majority	.31	.22	.31	.31	.31
Case B—Identical regression lines, minority means smaller					
Minority	.10	.38	.10	.27	.32
Majority	.36	.31	.36	.34	.32
Case C—Different regression slopes, minority slope smaller					
Minority	.01	.47	.00	.25	.33
Majority	.43	.21	.44	.36	.33
Case D—Different regression slopes and intercepts, minority slope smaller					
Minority	.00	.14	.00	.18	.35
Majority	.40	.38	.40	.38	.35
Case E—Different regression slopes and intercepts, minority slope larger					
Minority	.08	.14	.09	.26	.32
Majority	.35	.34	.35	.33	.32

from that table, potentially successful minority applicants have almost no chance of selection under the regression and employer's models in Cases C and D. Their chances of selection are better under the Darlington model than for their cohorts in the majority group in Cases A, B, and C but poorer than majority applicants in Cases D and E. Only under the equal opportunity model are those chances the same in each group in each case.

The different models also yield selectees with different chances of success once in the institution or on the job. The proportion of selectees who will succeed is a figure of great importance to the selecting institution in that the institution seeks to minimize failures. An overall expected success rate in the selected group can be obtained by computing $Pr_i\{Y > Y_p \mid X > X_i\}$. By assuming bivariate normality of X and Y, the required probability can be computed using tables of the bivariate normal distribution. The expected success rates thus computed are reported in Table 5.

As can be seen from Table 5, although there are different expected success rates in the minority and majority groups, the overall expected success rates differ very little. The most difference occurs in Cases C and D although the drop from the employer's model to the equal opportunity model is only .88 to .81 in Case C and .81 to .75 in Case D.

TABLE 5

Expected Success Rates of Selectees

$Pr\{Y > Y_p \mid X > X_i\}$

	Regression model	Darlington model	Employer's model	Thorndike model	Equal opportunity model
Case A–Parallel regression lines, minority intercept larger					
Minority	.78	.68	.78	.78	.78
Majority	.78	.82	.78	.78	.78
Overall	.78	.76	.78	.78	.78
Case B–Identical regression lines, minority means smaller					
Minority	.70	.59	.70	.64	.61
Majority	.76	.78	.76	.78	.78
Overall	.76	.74	.76	.75	.75
Case C–Different regression slopes, minority slope smaller					
Minority	.71	.57	.77	.61	.60
Majority	.87	.92	.88	.89	.90
Overall	.87	.77	.88	.83	.81
Case D–Different regression slopes and intercepts, minority slope smaller					
Minority	.55	.42	.58	.40	.33
Majority	.81	.82	.81	.83	.83
Overall	.81	.80	.81	.80	.75
Case E–Different regression slopes and intercepts, minority slope larger					
Minority	.72	.66	.71	.56	.53
Majority	.71	.71	.71	.71	.71
Overall	.71	.71	.71	.70	.69

Discussion

The preceding examples bring into focus the differences in the idea of fairness used in each of the models examined. In addition, because the hypothetical cases are realistic imitations of many common selection situations, the examples show the actual different effects to be expected from using the different models in selection.

Both the quota model and the Darlington model use an explicit statement of value associated with the selection of members of some group over another group. Thus when social values dictate that some group be favored in the selection process, each of these models provides a means of implementing those values. While both models may be useful in some situations, their applications are limited by the a priori favoritism and the fact that they place other factors above the importance of the criterion. In cases in which selection involves cost and effort and then possible failure, the social good of selection must be carefully balanced against the possible negative effect of large-scale failure. In the quota model and Darlington model, in the concern with the value of selection the importance of the criterion may be overlooked.

In the regression model and employer's model, the concern is solely with the importance of the criterion at the expense of ideas of fairness to the applicant. Both models provide high expected success rates among the selectees and are therefore advantageous to the selecting institution, but the applicant may see little advantage in them. The potentially successful applicant's concern is that he have a fair chance of selection regardless of the group of which he is a member. In cases in which the regression line for his group is such that even though he could succeed he will have less chance of selection than members of other groups, no selecting institution's concern with selection of applicants with highest predicted criterion scores will receive much sympathy. And when poor prediction in his group is the cause of his chances of selection being lowered, the applicant will rightly blame the institution for its failure to find a good predictor—a situation for which the applicant should not be penalized.

In many situations the rights of the potentially successful applicants to fairness should be of primary concern. Under the equal opportunity model such applicants are guaranteed equal chance of selection regardless of group membership. This procedure places the burden of improving prediction on the selecting institution and even allows for the use of different predictors in different groups. If present tests tend to work poorly in predicting criterion scores for minority members, under the equal opportunity model the selecting institution must compensate for the poor predictors by selecting more, not fewer, minority students.

It should be noted that the equal opportunity model does not eliminate or de-emphasize the use of tests or other predictor variables in the selection process. In fact, under this model the selecting institution and the potentially successful applicant both benefit from the use of better tests or other predictors within each group. However, use of this model should provide an important step in the direction of insuring minority group members that their rights are not being subverted by tests chosen by and constructed by majority group members.

Conclusions

The selection procedure most commonly used is the regression model. That model was designed to meet the needs of the selecting institution to select successful persons. Where a large investment of money or time is made in each individual selected, the institution naturally seeks those most likely to succeed. As can be seen in the examples considered here, both the regression model and the employer's model directly meet these needs of the selecting institution.

However, it seems likely that the applicant's rights will be ruled dominant to those of the selecting institution by many people, especially when the investment is not prohibitive. Certainly in the field of education and other areas in which the social benefits of selection are considerable, fairness to the applicant is very important, especially when fairness and equal opportunity mean not the opportunity to fail but the opportunity to

succeed as under the equal opportunity model. When public institutions such as government agencies or public colleges and universities are the selecting institutions, it becomes even more difficult to argue that the institution's right to successful selectees exceeds the potentially successful applicant's right to the same chance of selection regardless of his group membership.

In some situations cultural-social values may be judged dominant to this right to equal opportunity. However, in the common selection situations encountered in empirical studies of bias in selection of racial/ethnic minority members as illustrated in Cases B through E, even the use of the equal opportunity model will have a decided social effect in providing for the selection of more minority group members than under the commonly used regression model.

The use of the equal opportunity model can assure all groups of an intuitively meaningful and defensible type of fairness. In addition, the model promotes the equal opportunity of minority group members who have missed selection under commonly used models of bias in many selection situations. At the same time the model does not dramatically reduce the expected success rate among those selected—a major concern of the selecting institution—in the common situations considered here. Therefore, it is concluded that the equal opportunity model should be widely implemented in many college and employment selection situations.

REFERENCES

Anastasi, A. *Psychological testing.* (3rd ed.) New York: Macmillan, 1968.

Bartlett, C. J., & O'Leary, B. S. A differential prediction model to moderate the effects of heterogeneous groups in personnel selection and classification. *Personnel Psychology*, 1969, 22, 1-17.

Bowers, J. The comparison of GPA regression equations for regularly admitted and disadvantaged freshmen at the University of Illinois. *Journal of Educational Measurement*, 1970, 7, 219-225.

Cleary, T. A. Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement, 1968*, 5, 115-124.

Darlington, R. B. Another look at "culture fairness." *Journal of Educational Measurement*, 1971, 8, 71-82.

Einhorn, H. J., & Bass, A. R. Methodology considerations relevant to discrimination in employment testing. *Psychological Bulletin*, 1971, 75, 261-269.

Guion, R. Employment tests and discriminatory hiring. *Industrial Relations*, 1966, 5, 20-37.

National Bureau of Standards. *Tables of the bivariate normal distribution function and related functions*. National Bureau of Standards Applied Mathematics Series 50, 1959.

Temp, G. Validity of the SAT for blacks and whites in thirteen integrated institutions. *Journal of Educational Measurement*, 1971, 8, 245-251.

Thorndike, R. L. Concepts of culture-fairness. *Journal of Educational Measurement*, 1971, 8, 63-70.

ACT Research Reports

This report is Number 51 in a series published by the Research and Development Division of The American College Testing Program. The first 26 research reports have been deposited with the American Documentation Institute, ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington, D.C. 20540. Photocopies and 35 mm. microfilms are available at cost from ADI; order by ADI Document number. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress. Beginning with Research Report No. 27, the reports have been deposited with the National Auxiliary Publications Service of the American Society for Information Science (NAPS), c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001. Photocopies and 35 mm. microfilms are available at cost from NAPS. Order by NAPS Document number. Advance payment is required. Printed copies ($1.00) may be obtained, if available, from the Publication and Information Services Division, The American College Testing Program, P.O. Box 168, Iowa City, Iowa 52240. A check or money order must accompany the request.

The reports since May 1970 in this series are listed below. A complete list of the reports can be obtained by writing to the Publication and Information Services Division, The American College Testing Program, P. O. Box 168, Iowa City, Iowa 52240.

No. 34 *Research Strategies in Studying College Impact*, by K. A. Feldman (NAPS No. 01211; photo, $5.00; microfilm, $2.00)

No. 35 *An Analysis of Spatial Configuration and Its Application to Research in Higher Education*, by N. S. Cole, & J. W. L. Cole (NAPS No. 01212; photo, $5.00; microfilm, $2.00)

No. 36 *Influence of Financial Need on the Vocational Development of College Students*, by A. R. Vander Well (NAPS No. 01440; photo, $5.20; microfilm, $2.00)

No. 37 *Practices and Outcomes of Vocational-Technical Education in Technical and Community Colleges*, by T. G. Gartland, & J. F. Carmody (NAPS No. 01441; photo, $6.80; microfilm, $2.00)

No. 38 *Bayesian Considerations in Educational Information Systems*, by M. R. Novick (NAPS No. 01442; photo, $5.00; microfilm, $2.00)

No. 39 *Interactive Effects of Achievement Orientation and Teaching Style on Academic Achievement*, by G. Domino (NAPS No. 01443; photo, $5.00; microfilm, $2.00)

No. 40 *An Analysis of the Structure of Vocational Interests*, by N. S. Cole & G. R. Hanson (NAPS No. 01444; photo, $5.00; microfilm, $2.00)

No. 41 *How Do Community College Transfer and Occupational Students Differ?* by E. J. Brue, H. B. Engen, & E. J. Maxey (NAPS No. 01445; photo, $5.50; microfilm, $2.00)

No. 42 *Applications of Bayesian Methods to the Prediction of Educational Performance*, by M. R. Novick, P. H. Jackson, D. T. Thayer, & N. S. Cole (NAPS No. 01544; photo, $5.00; microfilm, $2.00)

No. 43 *Toward More Equitable Distribution of College Student Aid Funds: Problems in Assessing Student Financial Need*, by M. D. Orwig (NAPS No. 01543; photo, $5.00; microfilm, $2.00)

No. 44 *Converting Test Data to Counseling Information*, by D. J. Prediger (NAPS No. 01776; photo, $5.00; microfiche, $2.00)

No. 45 *The Accuracy of Self-Report Information Collected on the ACT Test Battery: High School Grades and Items of Nonacademic Achievement*, by E. J. Maxey, & V. J. Ormsby (NAPS No. 01777; photo, $5.00; microfiche, $2.00)

No. 46 *Correlates of Student Interest in Social Issues*, by R. H. Fenske, & J. F. Carmody (NAPS No. 01778; photo, $5.00; microfiche, $2.00)

No. 47 *The Impact of College on Students' Competence to Function in a Learning Society*, by M. H. Walizer, & R. E. Herriott (NAPS No. 01779; photo, $5.00; microfiche, $2.00)

No. 48 *Enrollment Projection Models for Institutional Planning*, by M. D. Orwig, P. K. Jones, & O. T. Lenning (NAPS No. 01780; photo, $5.00; microfiche, $2.00)

No. 49 *On Measuring the Vocational Interests of Women*, by N. S. Cole (NAPS No. not available at this time.)

No. 50 *Stages in the Development of a Black Identity*, by W. S. Hall, R. Freedle, & W. E. Cross, Jr. (NAPS No. not available at this time.)